



CD-doc-666-v2.doc
GCC Automatic Shutdown Plan

David J. Ritchie
Operations, Computing Division

February 2, 2005

Introduction

The purpose of this document is to describe the plan for providing an automatic shutdown capability for the analysis computing facilities in GCC.

Background

GCC, the Grid Computing Center, sited within the building formerly known as "Wideband," will house a large number of PCs (up to 2800) performing high energy physics computations and connected to the laboratory network through a number (~5) of subnets. The PCs, which generally run Linux, are powered via a 1000 KVA UPS system which obtains *its* power from the laboratory electrical feeders supplied by Commonwealth Edison via the laboratory substation. Cooling is provided through a number (~8) of Liebert Computing Room Air Conditioners (CRACs).

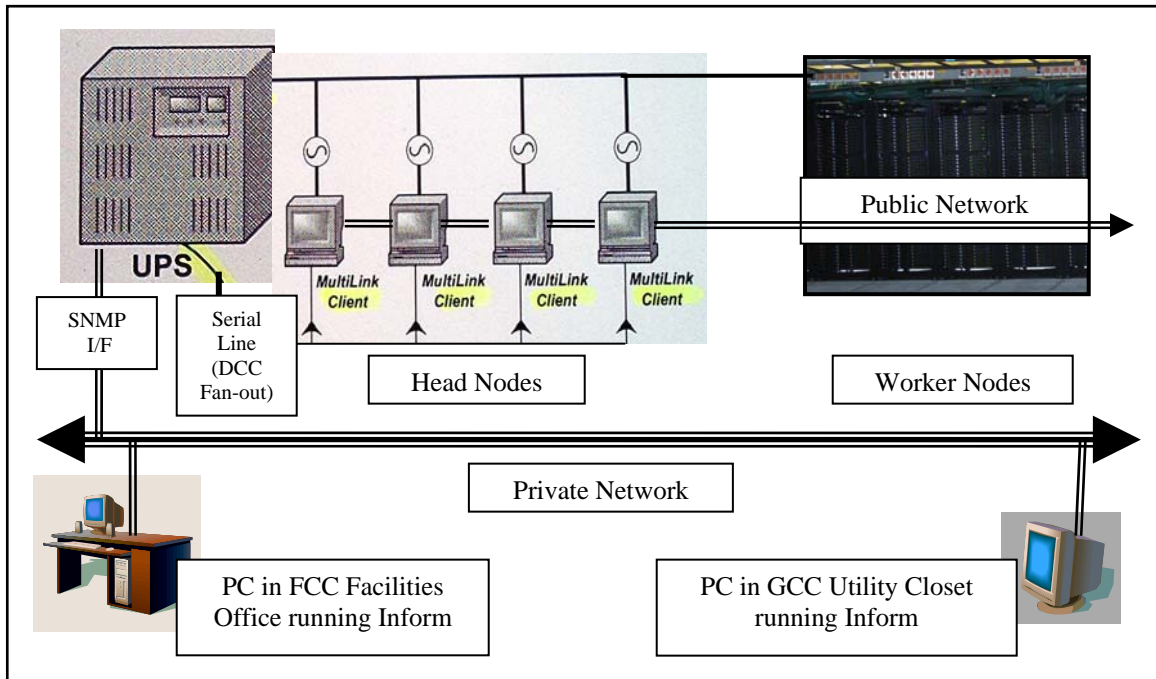
The purpose of the UPS is to make it possible for the PC's to shutdown in a controlled way (i.e., "a soft landing") in the event of a failure of the laboratory power. Experience has shown that by providing such a soft landing capability, the occurrence of hardware problems due to unexpected power outages and the cost to fix the problems will be minimized and possibly even avoided altogether.

The capacity of the UPS makes it possible for a totally installed GCC (i.e., 2800 PCs) to have power for a maximum of 13 minutes. The intent, of course, is to have the PCs shutdown in a much shorter time – say, two to five minutes.

Note that the CRACs are not powered through the UPS but instead use the usual laboratory power. It is not a goal to keep the CRACs powered during the soft shutdown process – instead, it is intended to shut the PCs down before the temperature rises to a point where it would be dangerous to the equipment.

Plan for Implementing Automatic Shutdown at GCC

The plan is to purchase the MultiLink capability as diagrammed below:



Hardware Components and Details...

1. UPS in GCC with *both* Direct Contact Closure *and* SNMP I/F card
2. Serial Line connecting UPS to Head Nodes via Direct Contact Closure “fan-out” (~\$758)
3. Windows PC running Inform located in GCC Utility Closet to provide on-location access to UPS and CRACs status (~\$1000 for PC and \$1500 for Inform License)
4. Linux Head Node PCs running the MultiLink v1.5 Client; supplied by CD Groups Responsible for Farms; (~\$250 MultiLink License for up to 10 Clients); located in GCC Computing Room
5. Windows PC running Inform located in FCC Facilities Office to provide remote status of UPS and CRACs; supplied by Facilities Office; (~\$1000 for PC; ~\$1500 for Inform License)
6. Private Fiber Network between FCC Facilities Office and GCC; runs SNMP protocol in a closed environment.

Software Components and Details...

1. Inform on Windows PC located in GCC Utility Closet. This monitors via the private network and the SNMP I/F the CRACs and UPS—e.g., UPS events, such as:
 - a. on-battery,
 - b. low battery,
 - c. return to normal (i.e., on utility power), and
 - d. weak battery
2. MultiLink 1.5 in each Client Head Node
 - a. The MultiLink 1.5 client receives “on-battery” indications via serial line; starts a timer that runs for intervals specified in the MultiLink 1.5 configuration file on the Head Node.
 - b. Provided that the started timer expires without being canceled by a “return to normal” event, the client executes an “on-battery” script and later an “os-shutdown” script.
 - c. These scripts, which it is imagined are written by, and are the responsibility of, the CD Groups supporting the associated worker nodes, cause the analysis programs running in the worker nodes to stop (“on-battery” script) and cause the os and the power controllers to shutdown (“os-shutdown” script).
3. MultiLink Network Shutdown License (~\$250.00)
4. Inform License – supports local PC in GCC Utility closet and remote PC located in FCC Facilities Office (~\$3000.00)
5. Software on CDROM (~\$50.00)

Cost Summary

Item	Amount
SNMP I/F Card for UPS (est.)	\$500.00
Serial Line and Fan-out for Direct Contact Closure	\$758.00
GCC Windows PC running Inform (est.)	\$2500.00
One MultiLink™ 3.5 Advanced Shutdown and MultiLink™ 1.5 Shutdown software. (Provided on CDROM).	50.00
One MultiLink™ network shutdown license for up to 10 computers. (Provided on CDROM).	\$250.00
Windows PC running Inform located in FCC Facilities Office (est.)	\$2500.00
Total	\$6558.00

Advantages Summary

1. The UPS is well monitored using manufacturer supported instrumentation.
2. There are two alert points so that implementers of head node scripts have some options as to whether to trigger on the on-batteries event or on the os-shutdown event. This makes it possible, for example, to interpret the first as directed towards the analysis programs while the second as directed towards the power controllers.
3. There is a monitoring method—namely the utility room PC—at GCC which is more capable and more readable than the front panels for the UPS and the CRACs.
4. There is a remote monitoring method—namely the PC in the Facilities Office at FCC which provides a view of the state of things at GCC from afar. It is done over a private network so that the software's use of the SNMP protocol is not exposed.

Extensibility Summary

The extensibility of this automatic shutdown solution allows:

1. Extension of the same system to additional head nodes (i.e., extension to 100 PCs would cost \$7.50/PC or \$750.00 for the Network Shutdown software licenses) is possible.

It is noted that it would be necessary to purchase fan-out hardware and to run a serial line to each of the PC's. This would become cumbersome for hundred's of PC's but it is the price one has to pay to avoid exposing SNMP protocol to the public network.

2. Extension of the same system to monitor the Liebert Computer Room Air Conditioners within GCC is possible.
3. Extension of the same system to monitor from FCC additional UPS systems either within GCC or at LCC or elsewhere is also possible.

Difficulties Overcome

The principal difficulty overcome in this plan is that it is not necessary to expose the use of SNMPv1 as a communication protocol to the public network. That protocol (particularly at v1) is not felt to be very secure.

It may be that there are different mechanisms for "hardening" the MultiLink use of SNMPv1. There have been discussions concerning having the shutdown scripts validate their requested initiation. However, the software employs SNMPv1 in other ways so this may not be sufficient. We are open to suggestions in this area.

Alternatives Considered

We have considered a number of alternatives to the above automatic shutdown plan (Thanks to Keith Chadwick for bringing these to our attention):

1. Netbotz—(<http://www.netbotz.com>) We have looked at the Netbotz management and monitoring device option. While this option has transmission of alerts via SSL and so avoids the SNMPv1 problem, it does not have the capability as near as I can tell of executing a script on the Head Node when an alert occurs.

2. Omnitronix—(<http://www.omnitronix.com>) We have examined the information on the Omnitronix web site. Its offerings make use of SNMP extensively. In particular, it does have AlarmManager software but it only apparently on the Windows operating system. We need the ability to send alarms to Linux boxes, however, so this is ruled out.